CDF/PUB/TOP/PUBLIC/10395

Version 1.0

Search for Heavy Top $t' \to Wq$ in Lepton Plus Jets Events in $\int \mathcal{L}dt = 5.6$ fb^{-1}

John Conway, David Cox, Robin Erbacher
Will Johnson, Tom Schwarz
University of California, Davis

Andrew Ivanov
Kansas State University

Alison Lister
University of Geneva

March 8, 2011

Abstract

We search for pair production of the heavy top (t') quarks pair decaying to Wb final states using a 5.6 fb^{-1} data sample.

We reconstruct the mass of the t' quark (M_{rec}) and perform a two dimensional-fit of the observed (H_T, M_{rec}) distribution to discriminate the new physics signal from Standard Model backgrounds. We exclude Standard Model fourth-generation t' quark with mass below 358 GeV at 95%CL.

Contents

1 Introduction

The top quark is a relatively recent addition to the array of particles that can be produced in the laboratory. Since its discovery the top quark data collected at the Tevatron have been an active testing ground for the validity of the Standard Model (SM). The top quark is unique because of its large mass near 173 GeV, which distinguishes it from the other fermions of the SM and is similar to the masses of the weak force carriers (W and Z) and the expected mass range for the proposed Higgs boson.

Because of the large top quark mass, the top quark final decay products are very energetic. The leptons and jets from top decays have on average higher transverse momenta as compared to those produced from other SM processes. While these kinematic features are often employed to discriminate the top quark signal from SM backgrounds, there is a number of new physics models predicting heavy quarks with masses above the one of the top quark and producing event signatures similar to those from top quark decays. The simplest extension of the SM with three generations is a fourth chiral generation of massive fermions. The fourth generation is predicted in a number of theories [1, 2], and although historically have been considered disfavored, is in a good agreement with electroweak precision data [3, 4].

To avoid $Z \to \nu\bar{\nu}$ constraint from LEP I a fourth generation neutrino ν_4 must be heavy: $m(\nu_4) \gtrsim m_Z/2$, where m_Z is the mass of Z boson, and to avoid LEP II bounds a fourth generation charged lepton ℓ_4 must have $m(\ell_4) \gtrsim 101$ GeV. At the same time due to sizeable radiative corrections masses of fourth generation fermions cannot be much higher the current lower bounds and masses of new heavy quarks t' and b' should be in the range of a few hundreds GeV [4], that could be accessible at Tevatron collider. In addition, a small mass splitting between t' and b' is preferred, such that $m(b') + m(W) > m(t')$, and t' decays predominantly to Wq (a W boson and a down-type quark $q = d, s, b$) [4, 5].

In the four-generation model the present bounds on the Higgs are relaxed: the Higgs mass could be as large as 500 GeV [4, 5], which could resolve the conflict between the SM prediction for the Higgs mass and the LEP II direct lower limit [7]. Furthermore, the CP violation is significantly enhanced to the magnitude that might account for the baryon asymmetry in the Universe [8]. Additional chiral fermion families can also be accommodated in supersymmetric two-Higgs-doublet extensions of the SM with equivalent effect on the precision fit to the Higgs mass [9].

Another possibility is heavy exotic quarks with vector couplings to the W boson. Contributions to radiative corrections from such quarks with mass M decouple as $1/M^2$ and easily evade all experimental constraints. For example, the "beautiful mirrors" model [10] motivated by 3σ discrepancy between the hadronic and leptonic asymmetry measurements from LEP II, which result in controversial predictions for the Higgs mass [7, 11], solves the problem by introducing a new vectorlike fermion doublet, a mirror copy of the standard quark doublets with a heavier version of the SM top decaying to Wb.

A heavy top-like quark also appears in Little Higgs models [12], which evade the hierarchy problem by introducing a minimal set of gauge and fermion fields in the context of a large-extra-dimension framework. In particular, models in which T-parity is conserved suggest a massive top-like quark which can decay to Wq and have a mass of approximately 500 GeV [13].

Thus, there is a number of well-motivated scenarios predicting a heavy top-like particle decaying into a W boson and a bottom quark. In this work we search for pair production of such hypothetical new quarks using events characterized by a high-p_T lepton, large $\not\!\!E_T$, and

multiple hadronic jets. We refer to the hypothetical new quark as t'. We assume that the new quark is heavier than the top, and for the purpose of setting limits we assume that the new heavy quarks are produced strongly and have the same couplings as the SM quarks of the first three generations.

The previous iteration of this analysis is documented in [14]. This analysis uses a larger dataset of 5.6 fb^{-1} and assumes the t' quark decays specifically to a bottom quark allowing the usage of b-tagging for the first time.

Namely, with respect to the previous iteration of the analysis:

- We add additional data
- We require a b-tag in selection.
- We add additional systematics assocaited with b-tagging.

More details on each of these items is described further in the note.

2 Event Selection

The CDF detector is described in detail in [6]. We use a data sample corresponding to 5.6 fb^{-1} of integrated luminosity.

The following event selection criteria are applied

- one and only one high-p_T ($p_T \geq 20$ GeV/c) isolated electron or muon
- large missing transverse energy (≥ 20 GeV)
- at least four energetic jets ($E_T \geq 20$ GeV)
- at least one tagged b-jet

The dominant contributing backgrounds after these cuts are from top pair production as well as W+jets. Much smaller backgrounds include QCD, electroweak processes, diboson and single top production and Z + jets. All of these processes except for QCD are modeled using MC simulation.

Using these event selection criteria we observe a total of 1441 events.

3 Analysis Method

The variable

$$H_T = \sum_{jets} E_T + E_{T,\ell} + \not\!\!E_T, \tag{1}$$

serves as a good discriminator between Standard Model and new physics processes associated with production of high mass particles.

In addition, we make use of the fact that t' decay chain is identical to the one of the top quark, and reconstruct its mass similarly to as it is done in the top quark mass measurement analyses. We adopt the template method for the top quark mass reconstruction, which is based on the χ^2-fit of kinematic properties of final top decay products.

We consider only 4 highest E_T jets in the mass reconstruction. For each event there are total $4!/2 = 12$ combinations of assigning 4 jets to partons. In addition, there are two solutions for unknown P_z neutrino momentum. The MINUIT minimization is performed for each of the 24 combinations, and then the permutation with the lowest value of χ^2 is selected. The χ^2 is given by the following expression:

$$\chi^2 = \sum_{i=\ell,4jets} \frac{(p_T^{i,fit} - p_T^{i,meas})^2}{\sigma_i^2} + \sum_{j=x,y} \frac{(p_j^{UE,fit} - p_j^{UE,meas})^2}{\sigma_j^2}$$

$$+ \frac{(m_{jj} - m_W)^2}{\Gamma_W^2} + \frac{(m_{\ell\nu} - m_W)^2}{\Gamma_W^2} + \frac{(m_{bjj} - m_t)^2}{\Gamma_t^2} + \frac{(m_{b\ell\nu} - m_t)^2}{\Gamma_t^2}, \quad (2)$$

where invariant masses of W decay products m_{jj} and $m_{\ell\nu}$ are constrained to the pole mass of the W boson m_W, and masses of top and anti-top (t' and $\bar{t'}$) quarks are required to be equal. Jet, lepton and underlying event energies are allowed to float within their uncertainties, while the transverse component of neutrino momentum is calculated at each step of the fit, as follows

$$\vec{p}_T^{\,\nu} = -(\vec{p}_T^{\,\ell} + \sum \vec{p}_T^{\,jet} + \vec{p}_T^{\,UE}). \quad (3)$$

The longitudinal component p_z^ν is an unconstrained parameter in the fit and initialized with the value such that $m_{\ell\nu}$ acquires W pole mass m_W.

The m_t is the free parameter initialized with $m_t = 175$ GeV, and its value in the best fit is declared to be the reconstructed mass M_{rec} of top (or t' respectively).

Since in this analysis we are looking for $t' \to Wb$, b-tagging information is used in the mass reconstruction.

Unlike in top mass measurements we do not reject events that have a poor χ^2 for reconstructed events, but instead split events based on a good or bad χ^2 into separate categories.

In order to improve the discrimination power of our method and improve the sensitivity to a potential t' signal, we split the templates into four regions, based on the number of jets: $= 4$ or ≥ 5, and good or poor χ^2: $\chi^2 < 8$ and $\chi^2 > 8$. The sample of exactly 4 jets and a good χ^2 has the largest statistics due to the fact that majority of $t\bar{t}$ events (61% out of all ≥ 4 jets events) fall into this category. This is the region where the t' reconstruction performs the best, and in case of the t' signal in the data, one would hope to see the mass bump in the M_{rec} distribution.

Next, 18% of $t\bar{t}$ events fall into ≥ 5 jets, $\chi^2 < 8$ category. Here due to the additional jet the reconstruction performance is a bit worse. Smaller fractions of $t\bar{t}$ events fall into $\chi^2 > 8$ region, 14% and 7% of events for 4 and 5-jet bins respectively. The t' mass reconstruction is rather poor in these categories of events. However, because $t'\bar{t'}$ events are distributed more uniformly among all four categories of events (see Table 1), those are important to keep in order to increase acceptance to potential t' signal.

	$= 4$ jets	≥ 5 jets
$\chi^2 < 8$	31%	15%
$\chi^2 > 8$	29%	23%

Table 1: Fractions of t' events with mass of 400 GeV in four categories of events.

Thus we perform the search for the t' signal by employing a binned likelihood fit in both H_T and M_{rec} simultaneously for four different sets of templates:

- $= 4$ jet bin, $\chi^2 < 8$
- $= 4$ jet bin, $\chi^2 > 8$
- ≥ 5 jet bin, $\chi^2 < 8$
- ≥ 5 jet bin, $\chi^2 > 8$

The likelihood is defined as the product of the Poisson probabilities for observing n_i events in the bin i of (H_T, M_{rec}):

$$\mathcal{L}(\sigma_{t'}|n_i) = \prod_i P(n_i|\mu_i) \quad . \tag{4}$$

The expected number of events in each bin, μ_i, is given by the sum over all sources indexed by j, which is summed over all lepton categories:

$$\mu_i = \sum_j L_j \sigma_j \epsilon_{ij} \quad . \tag{5}$$

Here the L_j are the integrated luminosities, the σ_j are the cross sections, and the ϵ_{ij} are the efficiencies per bin of (H_T, M_{rec}).

We calculate the likelihood as a function of the $t'\bar{t}'$ cross section, and use Bayes' Theorem to convert it into a posterior density in $\sigma_{t'\bar{t}'}$. We can then use this posterior density to set an upper limit or measure the production rate of $t'\bar{t}'$.

The production rates for W+jets in the 4-jet bin and in the ≥ 5 jet bins are two free unconstrained independent parameters in the fit. Other parameters, such as the $t\bar{t}$ production cross section, lepton ID data/MC scale factors, integrated luminosity are related to systematic errors and treated in the likelihood as nuisance parameters constrained within their expected distributions. We adopt the profiling method for dealing with these parameters, i.e. the likelihood is maximized with respect to the nuisance parameters.

Taking this into account the likelihood takes the following expression:

$$\mathcal{L}(\sigma_{t'}|n_i) = \prod_{i,k} P(n_i|\mu_i) \times G(\nu_k|\tilde{\nu}_k, \sigma_{\nu_k}) \times \prod_{e,j} P(n_i|\mu_i) \times f_X(\nu_j|\tilde{\nu}_k, \sigma_{\nu_k}) \tag{6}$$

where ν_k are the nuisance parameters used in the morphing parameters (constrained by gaussian terms to their expectation) and ν_j are the nuissance parameters used in non-morphing parameters (constrained by log normal terms to their expectations), such as $\sigma_{t\bar{t}}$, L_j and etc. $\tilde{\nu}_k, j$ are their central nominal values and $\sigma_{\nu_k,j}$ are their uncertainties.

4 Systematic Uncertainties

4.1 Jet Energy Scale

The sensitivity to t' depends on knowing accurately the distribution of (H_T, M_{rec}). Therefore the largest source of uncertainty comes from the factor that has the greatest effect on the shape

of the kinematic distribution, which is due to the jet energy scale. Jets in the data and Monte Carlo are corrected for various effects, leaving some residual uncertainty.

This uncertainty results in possible shift in the H_T and M_{rec} distributions for both new physics and Standard Model templates. We take this effect into account by generating templates with energies of all jets shifted upward by one standard deviation (+1 templates) and downward (-1 templates) respectively.

Then we interpolate and extrapolate the expectation value μ_i at each bin i as follows:

$$\mu_i = \mu_{0,i} + \nu_{JES} \cdot (\mu_{+1,i} - \mu_{-1,i})/2 \tag{7}$$

where $\mu_{0,i}$ is the nominal expectation value, $\mu_{-1,i}$ and $\mu_{+1,i}$ are the expectation values from (-1) and (+1) templates respectively, and ν_{JES} is the nuisance parameter representing a relative shift in jet energy scale:

$$\nu_{JES} = \frac{\Delta_{JES}}{\sigma_{JES}} \quad . \tag{8}$$

It enters the likelihood (6) as a gaussian constraint penalty term: $G(\nu_{JES}|0,1) = \frac{1}{\sqrt{2\pi}} e^{-\nu_{JES}^2/2}$. This treatment of the systematic uncertainty in the likelihood is called vertical template morphing method.

4.2 W+jets Q^2 Scale

The effect of the choice of the appropriate Q^2 scale for W+jets production is evaluated by using the $W+$ jets Monte Carlo samples generated with different Q^2 settings. We make use of MC samples generated with a half and double of the nominal Q^2.

The Q^2 systematic is then incorporated into the likelihood in a manner similar to the Jet Energy Scale systematics, with an exception that it is only applied to $W+$ jets template. The expectation of $W+$ jets contribution in the bin i is given by

$$\mu_i = \mu_{Q^2=1.0,i} + \nu_{Q^2} \cdot (\mu_{Q^2=2.0,i} - \mu_{Q^2=0.5,i})/2, \tag{9}$$

where the parameter ν_{Q^2} is gaussian constrained in the likelihood.

4.3 Initial and Final State Radiation

We make use of $t\bar{t}$ samples that simulate the effect of increasing and decreasing the initial and final state radiation in $t\bar{t}$ events. The shifted templates ("IFSR less " and "IFSR more") serve as +1 and -1 σ templates and are incorporated into the likelihood as the Jet Energy Scale and the $W+$ jets Q^2 systematics. The morphing is only performed for $t\bar{t}$ template. In principle, the initial and final state radiation also effects the shapes of the t' H_T and M_{rec} distribution. However, the effect of this shift is tiny for t'. It changes the mean of the distribution by 2%. On the contrary, the effect from $t\bar{t}$ is non-negligible because it is a large background.

4.4 Integrated Luminosity

The integrated luminosity uncertainty is taken to be 5.8, and represented by an additional constrained parameter multiplying all contributions except for the QCD background and $W+$ jets, which are normalized independently.

4.5 Lepton ID

We make use of the lepton ID scale factors and trigger efficiencies that are standard for CDF analysis, and apply it to MC-based backgrounds only, except for $W+$ jets, which floats independently. The uncertainty due to those is 1% and is applied in quadrature with the uncertainty due to the NLO theoretical cross sections.

4.6 PDF Uncertainty

The Parton Distribution Functions (PDFs) are not precisely known, and this uncertainty leads to a corresponding uncertainty in the predicted cross sections, as well as in the acceptance. The first is a major part of the NLO theoretical cross section. The latter is estimated to be 1% from the $t\bar{t}$ cross section analyses, and is summed in quadrature with the uncertainty due to theory.

4.7 Theory Uncertainty

The theory uncertainty in the t' cross section is about 10% (see Table 2), which is mainly due to uncertainty in PDFs ($\sim$ 7%). The other effect comes from uncertainty in the choice of the Q^2 scale.

We take the theory uncertainty in $t\bar{t}$ cross section fully correlated with the one of t', and introduce it into the likelihood as a single nuisance parameter: $\nu_{theory} = \nu_{theory}(m'_t)$, which is the same parameter used to constrain $t\bar{t}$ cross section to a theoretical value.

Cross sections for small electroweak backgrounds are also known with the precision of 10% and applied as independent nuisance parameters.

4.8 B-tag Uncertainty

The rate of b-tagging in top and t' is not perfectly modeled in MC and so we apply a b-tagging scale factor and take an uncertainty therein. We follow the standard joint physics scale factor of 0.95 with an uncertainty of 0.04. Using this uncertainty we generate altered templates at $\pm 1\sigma$ using which we apply a morphing uncertainty as is done with the JES uncertainty.

5 Results

We test sensitivity of our method by drawing pseudoexperiments from Standard Model distributions, i.e. assuming no t' contribution. Range of expected 95% CL upper limits with one standard deviation bandwidth is shown in Figure 1. The purple curve is the theory curve [15], the values of which are given in Table 2. The lower σ_{min} and upper σ_{max} limits are obtained using the CTEQ6M family of parton density functions with uncertainties, together with the study of the scale uncertainty [17].

$m(t')$ (GeV)	σ_{min} (pb)	σ_{center} (pb)	σ_{max} (pb)
180.0	4.9938	5.7476	6.2396
200.0	2.7815	3.1898	3.4525
220.0	1.5926	1.8236	1.9710
240.0	0.9299	1.0647	1.1515
260.0	0.5499	0.6302	0.6828
280.0	0.3281	0.3769	0.4096
300.0	0.1968	0.2268	0.2475
320.0	0.1183	0.1370	0.1502
340.0	0.0711	0.0828	0.0914
360.0	0.0426	0.0500	0.0555
380.0	0.0255	0.0301	0.0337
400.0	0.0152	0.0181	0.0204

Table 2: Theory values of t' cross section for given mass [15].

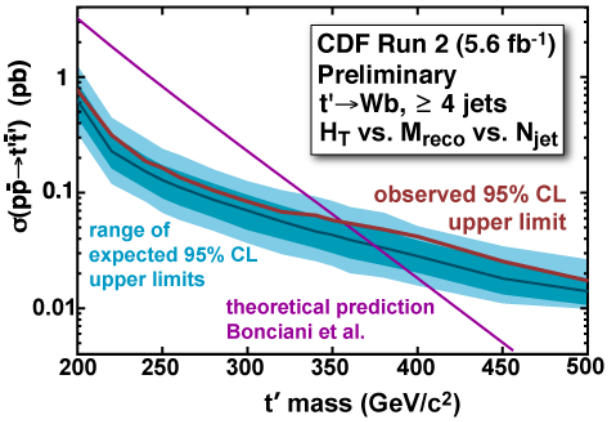

Figure 1: Limits.

From Figure 1 it follows that given no t' presence, this method is on average sensitive to setting an upper limit at 370 GeV t' mass.

We perform the analysis fit on the data and determine upper limits on the t' signal. The red curve in Figure 1 shows the final result, expressed as a 95% CL upper limit on the t' production rate as a function of t' mass. Table 3 shows the individual calculated limits along with expected limits from pseudo-experiments.

Distributions of H_T and M_{rec} in four different categories of events are shown in Figure 4. t' signal with mass of 360 GeV is normalized to the theoretical cross section value.

Based on these results we exclude at 95% CL the t' quark with mass below 358 GeV, given the true top mass is 172.5 GeV.

$m(t')$ (GeV)	expected limit (pb)	observed limit (pb)
180	$1.757\ ^{+0.729}_{-0.519}$	1.814
200	$0.563\ ^{+0.198}_{-0.178}$	0.581
220	$0.209\ ^{+0.099}_{-0.058}$	0.242
240	$0.142\ ^{+0.059}_{-0.041}$	0.139
250	$0.121\ ^{+0.047}_{-0.036}$	0.113
260	$0.104\ ^{+0.043}_{-0.029}$	0.106
280	$0.082\ ^{+0.034}_{-0.025}$	0.088
300	$0.065\ ^{+0.029}_{-0.018}$	0.076
320	$0.052\ ^{+0.023}_{-0.013}$	0.062
340	$0.044\ ^{+0.019}_{-0.011}$	0.057
350	$0.040\ ^{+0.019}_{-0.010}$	0.053
360	$0.037\ ^{+0.017}_{-0.010}$	0.054
380	$0.032\ ^{+0.013}_{-0.009}$	0.052
400	$0.028\ ^{+0.011}_{-0.008}$	0.049
450	$0.019\ ^{+0.007}_{-0.006}$	0.031
500	$0.013\ ^{+0.006}_{-0.003}$	0.020

Table 3: Expected and obtained limits on t' production cross section for given mass.

References

[1] J. Silva-Marcos JHEP 0212 (2002) 036; E. Arik, O. Cakir, S. A. Cetin and S. Sultansoy, Phys. Rev. D **66**, 033003 (2002), [arXiv:hep-ph/0204217]; E. Arik, O. Cakir, S. A. Cetin and S. Sultansoy, Acta Phys.Polon. B **37**, 2839 (2006), [arXiv:hep-ph/0502050].

[2] N. Borstnik *et al.*, Bled workshops in physics, Vol.7, No. 2, DMFA-Zaloznistvo, Ljubljana, Dec. 2006, [arXiv:hep-ph/0612250].

[3] V. A. Novikov, L. B. Okun, A. N. Rozanov and M. I. Vysotsky, Phys. Lett. B **529** (2002); V. A. Novikov, L. B. Okun, A. N. Rozanov and M. I. Vysotsky, JETP Lett. **76**, 127 (2002), [arXiv:hep-ph/0111028].

[4] G. D. Kribs, T. Plehn, M. Spannowsky and T. M. P. Tait, Phys. Rev. D **76**, 075016 (2007), [arXiv:0706.3718].

[5] P. H. Frampton, P. Q. Hung and M. Sher, Phys. Rept. **330**, 263 (2000), [arXiv:hep-ph/9903387].

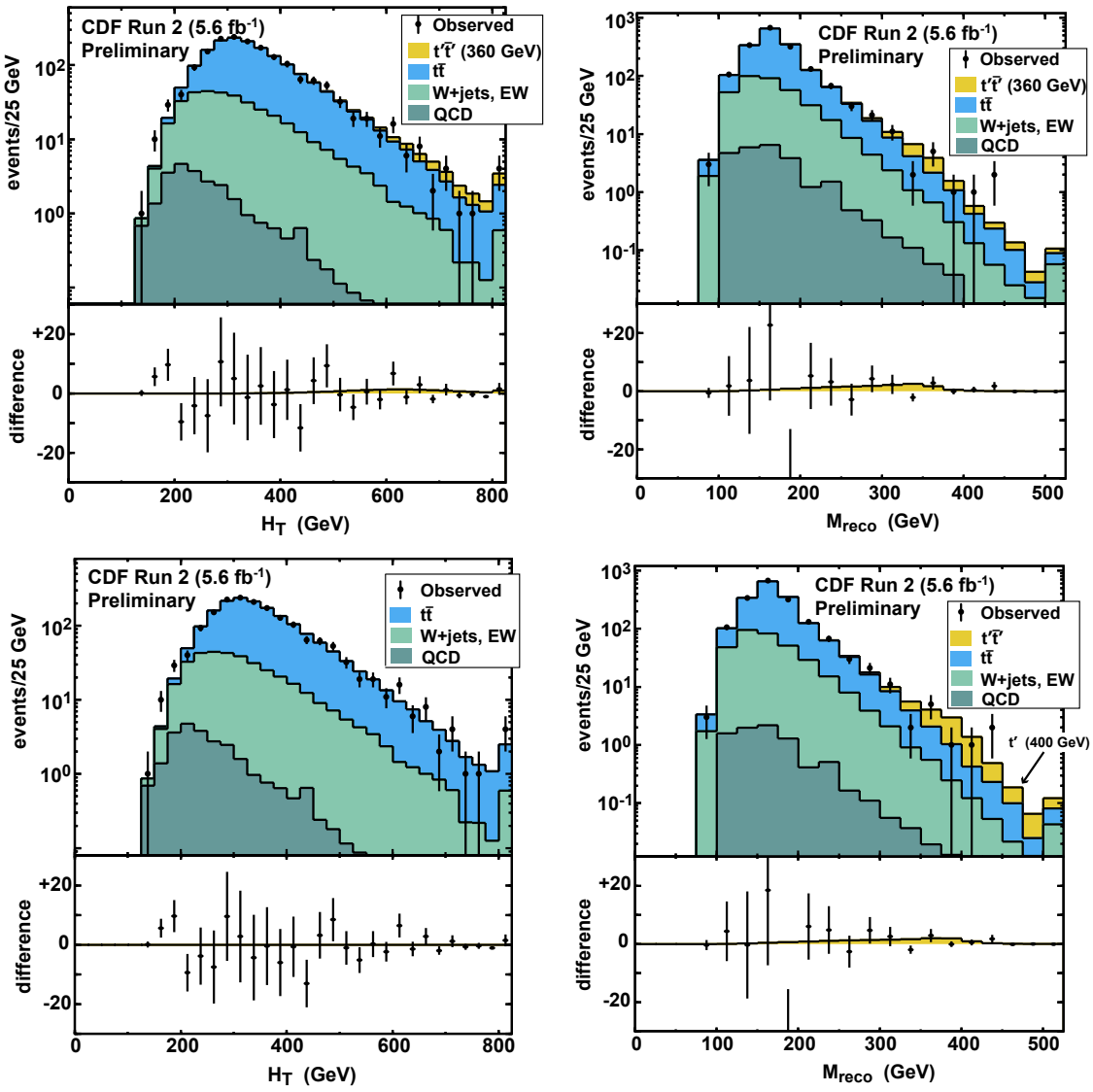

Figure 2: Log scale distributions of H_T and M_{rec} with $t'\bar{t}'$ signal of t' mass 400 GeV and no t' signal. All backgrounds set to their fitted results

[6] F. Abe, et al., Nucl. Instrum. Methods Phys. Res. A **271**, 387 (1988); D. Amidei, et al., Nucl. Instum. Methods Phys. Res. A **350**, 73 (1994); F. Abe, et al., Phys. Rev. D **52**, 4784 (1995); P. Azzi, et al., Nucl. Instrum. Methods Phys. Res. A **360**, 137 (1995); The CDFII Detector Technical Design Report, Fermilab-Pub-96/390-E

[7] C. Amsler *et al.* [Particle Data Group], Phys. Lett. B **667**, 1 (2008).

[8] W.-S. Hou, F. F. Lee, C.-Y. Ma, Phys. Rev. D **79**, 073002 (2009), [arXiv:0812.0064]; W.-S. Hou, [arXiv:0803.1234v3].

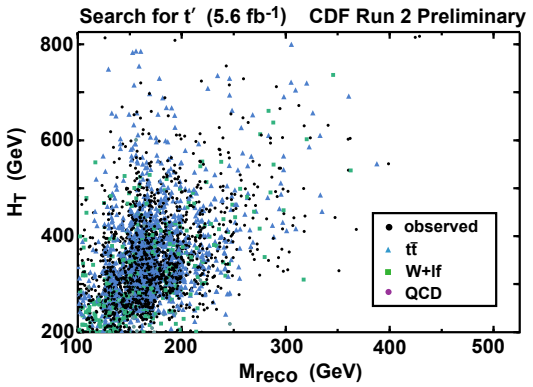

Figure 3: Scatter plot of events in H_T and M_{rec}

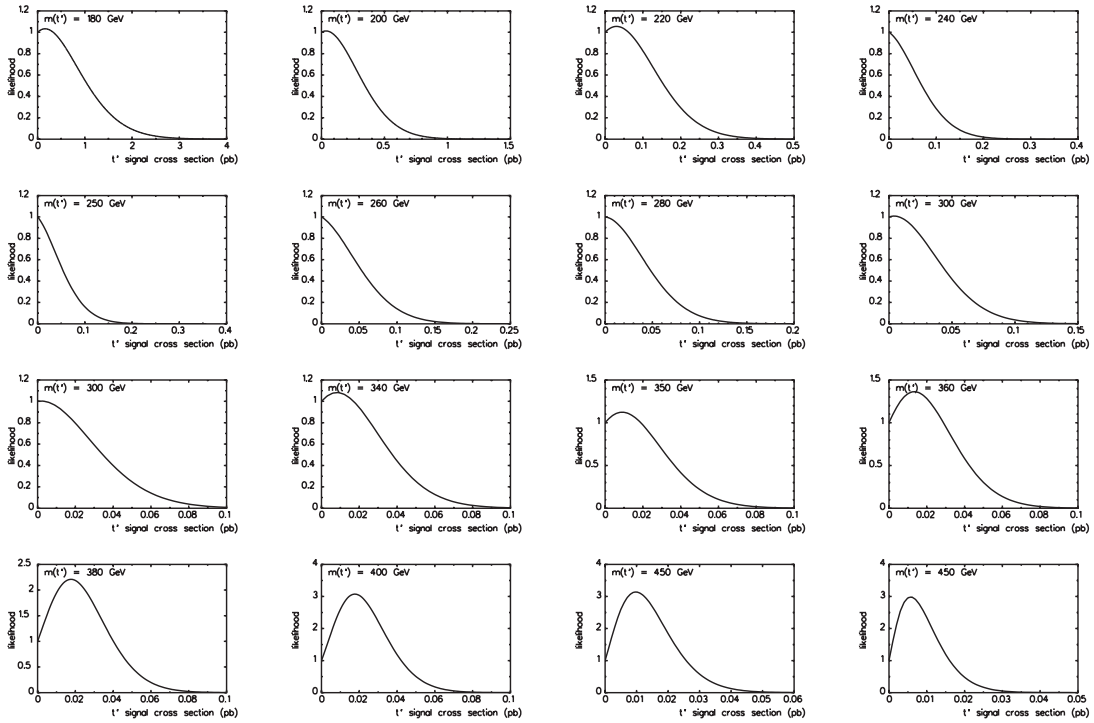

Figure 4: Likelihoods outputs from the fit

[9] H. J. He, N. Polonsky and S. Su, Phys. Rev. D **64**, 053004 (2001), [arXiv:hep-ph/0102144].

[10] D. E. Morrissey and C. E. M. Wagner, Phys. Rev. D **69**, 053001 (2004); D. Choudhury, T. Tait and C. E. M. Wagner, Phys. Rev. D **65**, 053002 (2002), [arXiv:hep-ph/0109097].

[11] M. Chanowitz, Phys. Rev. Lett. **87**, 231802 (2001), [arXiv:hep-ph/0104024].

[12] T. Han, H. Logan, B. McElrath, L.-T. Wang, Phys. Lett. B 563, 191 (2003); B. A. Dobrescu and C. T. Hill, Phys. Rev. Lett. 81, 2634 (1998).

[13] H. C. Cheng and I Low, J. High Energy Phys. **0408**, 61 (2004); D. E. Kaplan and M. Schmaltz, J. High Energy Phys. **566**, 375 (2006).

[14] The CDF Collaboration, CDF Note 10110

[15] M. Cacciari, S. Frixione, M. L. Mangano, P. Nason and G. Ridol, JHEP 0404 (2004) 068 [arXiv:hep-ph/0303085].

[16] J. Clark Cully, CDF Note 9628

[17] Private communication with M. L. Mangano.